DOCUMENT RESUME

ED 226 067 TM 830 114

AUTHOR Gottfredson, Gary D.
TITLE A Theory-Ridden Approach to Program Evaluation: A Method for Researcher-Implementer Collaboration. Report No. 330.
INSTITUTION Johns Hopkins Univ., Baltimore, Md. Center for Social Organization of Schools.
SPONS AGENCY National Inst. for Juvenile Justice and Delinquency Prevention (Dept. of Justice/LEAA), Washington, D.C.; National Inst. of Education (ED), Washington, DC.
PUB DATE Nov 82
GRANT 80-JN-AX-0005; NIE-G-80-0113
NOTE 27p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Delinquency; Educational Theories; *Evaluation Methods; Instructional Improvement; National Surveys; *Program Development; *Program Evaluation; Program Implementation; Researchers; *Research Methodology
IDENTIFIERS *Theory Based Evaluation

ABSTRACT

Social programs are often based on unclearly articulated theoretical rationales. The evaluations of these programs are also often relatively theory free, and therefore make limited contributions to theory. A Program Development Evaluation (PDE) approach, intended to integrate theory testing with the development and evaluation of action programs, is described and illustrated. This approach has been applied for over 2 years in the national evaluation of locally developed delinquency programs in 69 schools, and has resulted in (a) the implementation of several true experiments conducted in collaboration with project implementers, (b) an increase in the clarity of projects' theoretical rationales, and (c) the identification and measurement of theoretical intermediary variables as well as outcome variables. Based on the action research paradigm, the PDE approach appears to increase the theoretical and practical relevance of evaluation research. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED226067

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
  Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Center for Social Organization of Schools

Report No. 330
November 1982

A THEORY-RIDDEN APPROACH TO PROGRAM EVALUATION:
A METHOD FOR RESEARCHER-IMPLEMENTER COLLABORATION

Gary D. Gottfredson

The Johns Hopkins University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J Hollifield

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 830 114

STAFF

Edward L. McDill, Co-Director
James M. McPartland, Co-Director

Karl L. Alexander
Henry J. Becker
Jomills H. Braddock, II
Shirley Brown
Ruth H. Carter
Michael Cook
Robert L. Crain
Doris R. Entwisle
Joyce L. Epstein
James Fennessey
Samuel A. Gordon
Denise C. Gottfredson
Gary D. Gottfredson
Linda S. Gottfredson
Edward J. Harsch
John H. Hollifield
Barbara J. Hucksoll
Lois G. Hybl
Richard Joffe
Debbie Kalmus
Helene M. Kapinos
Nancy L. Karweit
Hazel G. Kennedy
Marshall B. Leavey
Nancy A. Madden
Kirk Nabors
Deborah K. Ogawa
Donald C. Rickert, Jr.
Laura Hersh Salganik
Robert E. Slavin
Jane St. John
Valerie Sunderland
Gail E. Thomas
William T. Trent
James Trone

A THEORY-RIDDEN APPROACH TO PROGRAM EVALUATION:

A METHOD FOR RESEARCHER-IMPLEMENTER COLLABORATION

Gary D. Gottfredson

Report No. 330

November 1982

This report is sponsored in part by Grant No. 80-JN-AX-0005 from the National Institute for Juvenile Justice and Delinquency Prevention, U.S. Department of Justice, and in part by Grant No. NIE-G-80-0113 from the National Institute of Education, U.S. Department of Education. The opinions expressed do not necessarily represent the position or policy of either agency. Published by the Center for Social Organization of Schools, supported in part as a research and development center by funds from the United States National Institute of Education, Department of Education.

Center for Social Organization of Schools
The Johns Hopkins University
3505 North Charles Street
Baltimore, Maryland 21218

Printed and assembled by the Centers for the Handicapped
Silver Spring, Maryland

The Center for Social Organization of Schools

The Center for Social Organization of Schools has two primary objectives: to develop a scientific knowledge of how schools affect their students, and to use this knowledge to develop better school practices and organization.

The Center works through five programs to achieve its objectives. The Studies in School Desegregation program applies the basic theories of social organization of schools to study the internal conditions of desegregated schools, the feasibility of alternative desegregation policies, and the interrelations of school desegregation with other equity issues such as housing and job desegregation. The School Organization program is currently concerned with authority-control structures, task structures, reward systems, and peer group processes in schools. It has produced a large-scale study of the effects of open schools, has developed Student Team Learning instructional processes for teaching various subjects in elementary and secondary schools, and has produced a computerized system for school-wide attendance monitoring. The School Processes and Career Development program is studying transitions from high school to post-secondary institutions and the role of schooling in the development of career plans and the actualization of labor market outcomes. The Studies in Delinquency and School Environments program is examining the role of school environments and experiences on delinquency.

The Center also supports a Fellowships in Education Research program that provides opportunities for talented young researchers to conduct and publish significant research, and to encourage the participation of women and minorities in research on education.

This report, prepared by the Delinquency and School Environments program, describes a method to make evaluation research more theoretically and practically useful.

Acknowledgements

An earlier version of this report was presented at the Meeting of the American Society of Criminology, Toronto, November 6, 1982. I am grateful for advice from Deborah Daniels, Denise C. Gottfredson, John Hollifield, and Jane St. John on a draft of this report, and for discussions with J. Douglas Grant and Carol Yamasaki.

A Theory-Ridden Approach to Program Evaluation:

A Method for Researcher-Implementer Collaboration

Abstract

Social programs are often based on unarticulated or unclearly articulated theoretical rationales. The evaluations of these programs are also often relatively theory free, and therefore make limited contributions to theory. A Program Development Evaluation (PDE) approach, intended to integrate theory testing with the development and evaluation of action programs, is described and illustrated. This approach has been applied for over two years in the national evaluation of locally developed delinquency programs in 69 schools, and has resulted in (a) the implementation of several true experiments conducted in collaboration with project implementers, (b) an increase in the clarity of projects' theoretical rationales, and (c) the identification and measurement of theoretical intermediary variables as well as outcome variables. Based on the action research paradigm, the PDE approach appears to increase the theoretical and practical relevance of evaluation research.

This report is about a critical task facing behavioral scientists--the task of designing research that advances both theory and practice. Educational and social programs are often based on unarticulated or unclearly articulated theoretical rationales. The evaluations of these programs are also often relatively theory free, and therefore make limited contributions to theory. My colleagues and I (Gottfredson, Note 1) have created a Program Development Evaluation (PDE) method intended to make possible the implementation and testing of stronger, more theoretically guided action programs. This approach has been applied for over two years in the national evaluation of locally developed delinquency prevention programs in 69 schools, and has resulted in (a) the implementation of several true experiments conducted in collaboration with project implementers, (b) an increase in the clarity of projects' theoretical rationales, and (c) the identification and measurement of theoretical intermediary variables as well as outcome variables. Based on the action research paradigm, the PDE approach appears to increase the theoretical and practical relevance of evaluation research.

The PDE method has evolved from efforts to help school systems and community-based organizations create effective programs and to evaluate those programs. A primary task for behavioral scientists is the development of knowlege, but researchers who have worked with practitioners struggling to solve problems know that the behavioral sciences have more to offer than the techniques of research, a smorgasboard of previously tested interventions, and a handfull of theories. Knowledge about the ways organizations behave and

about the psychology of decision making and planning in groups can be used to assist organizations not only to examine the results of their efforts, but also to increase organizational effectiveness in adopting theoretically plausible innovations and assuring their faithful implementation.

The PDE method provides a structure to merge the roles and activities of organization developer, theoretician, and evaluation researcher. This method is a form of action research (Chein, Cook, & Harding, 1948; Lewin, 1947; Sanford, 1970). The approach assumes that the prospects for promoting change are greatest when organizational decisionmakers' stake in the research is made clear by their own participation in the research. Decision makers and researchers collaborate through a continuing dialogue in which researchers provide feedback on the consequences of project action. Action research involves a cycle of hypothesis formulation and planning, action, evaluation and information feedback, and then renewed hypothesis formulation and planning. As the cycle is repeated, and information derived from project efforts and research is used in decision making, projects should become more effective--turning the process into an upward spiral of activity. And as the cycle is repeated, theory testing is refined, resulting in better, more practical theory.

Projects usually change over time on the basis of the experience gained as they develop (Wilkins & Gottfredson, Note 2). What Pearl (Note 3) has called "quality control" is needed to insure not just that a program is run according to the plan, but that a plan exists and is modified to coincide with the way a project, as it develops, is actually run. Many attempts to demonstrate the effectiveness of specifiable social programs have failed in part because plausible interventions were not implemented or their implementation was not documented (Quay, 1977; Sechrest, White, & Brown, 1979; Hall & Loucks, 1977), or the plan for the innovation was not clearly articulated at the outset (Sarason, 1971).

The PDE method facilitates the study of the development of a program by assisting in the planning process. It provides a mechanism by which an organization can make its plans and the theoretical rationale behind them explicit, and then engage in theory-relevant research as it goes about implementing and studying its activities. It also helps the researcher monitor and document project implementation as the project evolves. In short, the PDE method is an attempt to integrate theoretical research with evaluation and organizational development. Its action research approach to knowledge generation and organizational growth is derived from a tradition of concern for practical theory, useful research, and organizational change and development.

Guiding Principles

The evolution and practical use of the PDE method is guided by seven general principles. These principles are used to make decisions about the conduct of the evaluation research and to resolve ambiguities about appropriate next steps. Essentially, these guiding principles are a theory about doing theory-relevant research while increasing organizational effectiveness. Evaluation research conducted in accordance with these seven principles should

be most productive. Similarly, organizations that implement programs using these principles should be more effective that organizations that ignore them.

1. Action projects guided by explicit theories that can be translated into practice will be most effective. Not all theories can be translated into action by any given project, and all theories are not created equal. Science progresses by creating ideas and eventually rejecting the unproductive ones; some ideas have shown utility, and other things being equal, ideas that have survived previous empirical tests should have considerable precedence. The theory guiding a project should be a template for it, with decisions and interventions judged against it at the same time that the theory is judged by the effectiveness of the interventions. Projects will be implemented with most enthusiasm, be strongest, and contribute most to knowledge if (a) the theory is generated by or regarded as sensible by the project implementers themselves, and (b) the theory accords with evidence from previous research and evaluation.

2. Effective adoption of an innovation or other intervention is more likely when explicit plans for adoption are available, and perceived obstacles to organizational change are seen as likely to be overcome by a conceivable plan.

3. Effective implementation of an intervention or innovation is more likely if (a) blueprints for the intervention (manuals, protocols, etc.) are available, or if the intervention is structured by forms, rules, or operating procedures; and (b) implementation is subject to data guidance, i.e., observation and feedback of information to workers about the degree to which their behavior accords with the behavior specified by the blueprints for the intervention. Effective blueprints include plans for data guidance, and provide for documentation of the implementation of interventions.

4. Projects will increase in effectiveness under evaluation pressure. This pressure takes many forms, the most important of which are: (a) pressure to focus on theory in examining organizational behavior and the behavior of the organization's inhabitants; (b) pressure from potentially useful knowledge or information of relevance to the organization; (c) pressure from "personal knowledge" based on many sources, including direct observation or experience; (d) pressure from the rigorous, theory-based evaluation of intervention components; (e) pressure from the rigorous, theory-based evaluation of projects as a whole; (f) pressure from feedback about steps taken to adopt an innovation; and (g) pressure from feedback about steps taken to implement an intervention.

5. Projects internalizing these principles will behave in accordance with them more often than projects that simply comply with the application of them, and the former will therefore ultimately be more effective.

6. The more directly project implementers benefit from evaluation, the more evaluation will be integrated with project operations.

7. The interests of project implementers and evaluators coincide because one of the best ways to create communicable knowledge

is through the rigorous, theory-guided evaluation of well implemented interventions that can be described so well that others can understand what was done and therefore replicate them. Both implementers and researchers need rigorous evaluation, the adoption of innovations, well implemented interventions, thorough description, and theory.

These guiding principles appear to us to accord with sound organization development practice (French & Bell, 1978), sound field research practices (Empey, 1980), and practical wisdom in evaluation research (Tharp & Gallimore, Note 4). They were used to create the PDE method.

Organization Development as an Antecedent of PDE

One of the roots of Program Development Evaluation is the practice of organization development (OD). French and Bell (1978) characterize OD as a process involving action research that emphasizes normative change, is based in behavioral science, involves experience-based learning of intact work teams, and emphasizes goals and objectives. By characterizing OD as a process, French and Bell mean that OD is

> "not to be regarded as a one-shot solution to organizational problems, but more as a 'growing toward' greater effectiveness through a series of intervention activities over a period of time. . . . Changing the culture of . . . an entire organization is a long-term, involved process" (p. 69).

In addition, they see OD as a process involving rational, empirical strategies, but one that is even more dependent on normative-reeducative strategies: "The client defines what changes and improvements he or she wants to make, rather than the change agent; the change agent attempts to intervene in a mutual, collaborative way with the client as they together define problems and seek solutions; anything hindering effective problem solving is brought to light and publicly examined" (pp. 75-76). The emphasis on normative education is based on the assumption that behaviors are rooted in norms, values, or beliefs as well as in rationality and self-interest. OD is a data-based approach to planned change in which information is a spur to action. Unpleasant information is not to be avoided but rather treasured because it may lead to advancement, to clarification of problems. Typically, OD emphasizes concrete goal setting through the shared experience of a group in formulating plans. The on-the-job learning experience of an intact group is presumed to promote organizational and individual effectiveness.

The interactive, collaborative, participative approach, often used by behavioral scientists or OD specialists serving as consultants or facilitators of organizational planning and decision making, has much to offer in overcoming some of the difficulties a research effort may expect to face. First, increasing an organization's effectiveness should increase the likelihood that it will succeed in implementing interventions with a possibility of being shown to be effective when subjected to serious summative evaluation. Second, in the OD process, the scientist approaches an organization in a manner that may decrease the extent to which he or she is

perceived as an alien invader. By helping an organization clarify its goals and objectives, by assisting in creating open communication about problems, and by fostering the expectation that projects will change and develop over time, the researcher may come to be considered more as an insider, an entity to be trusted to convey useful news. And, the perspective that information, even uncomfortable information, is valuable in fostering growth and confronting important problems may decrease the organization's usual fear of evaluation. Finally, the links between OD and action research make the interjection of formal research possible.

The Program Development Evaluation method is in part a descendent of an OD method previously used by the Social Action Research Center (Blanton & Alley, Note 5) in a series of projects to manage and study social change. This predecessor, called the Program Development (PD) model, was developed through attempts to evaluate human service projects. In the PD model, feedback is a mechanism of project development that involves monitoring a project's environment, the implementation of strategies, and the achievement of goals. In practice, the Program Development specialist focuses on interaction with project implementers to assist in assessing needs, in articulating goals and more specific objectives, in analyzing a project's forcefield (environmental constraints and resources), and in developing strategies for change or implementation. Like other forms of OD, PD emphasizes participatory planning in part to foster normative-reeducation and in part to increase organizational and individual competencies in decision making and planning.

Action Research as an Antecedent of PDE

PDE has its roots in action research. According to French and Bell (1978), the origins of action research lie in the work of Dewey (1933), Collier (1945), and Lewin (1946). The roots of action research are, however, deeper than this. They can be traced back to the Baconian formulation of the scientific method, which specified three steps: (a) the formation of hypotheses, (b) the empirical testing of the hypotheses, and (c) the acceptance or rejection of the hypotheses (Deese, 1972). Action is taken to "twist the lion's tail" to learn about nature. Since Bacon, science has been active rather than speculative, historical, or reflective. Dewey translated the scientific method of problem solving for laypersons, and Collier and Lewin both applied the scientific method to solving practical social problems.

Collier (cited in French & Bell, 1978), a commissioner of Indian Affairs concerned with improving race relations, wrote of action research, claiming that:

> Research and then more research is essential to the program, that in the ethnic field research can be made a tool of action essential to all other tools, indeed that it ought to be the master tool. But we had in mind a particular kind of research, or, if you will, particular conditions. We had in mind research impelled from central areas of needed action. And since action is by nature not only specialized but also integrative . . . , our needed research must be

> of the integrative sort. Again, since the findings of the research must be criticized by them through their experience, the administrator and the layman must themselves participate creatively in the research, impelled as it is from their own area of need. (p. 94).

Broader attention was called to action research by Lewin (1947), an eminent and influential psychological theorist with a keen interest in the applications of psychology. He saw that cooperation between the change agent (or field worker) and the researcher is important for both planning and management:

> Planning starts usually with something like a general idea. For one reason or another it seems desirable to reach a certain objective. . . . The first step, then, is to examine the idea carefully in the light of the means available. Frequently more fact-finding about the situation is required. If the first period of planning is successful, two items emerge: an 'overall plan' of how to reach the objective and a decision in regard to the first step of the action. Usually this planning has also somewhat modified the original idea. The next period is devoted to executing the first step of the overall plan . . . (and) by certain fact-findings. . . . This . . . fact-finding has four functions. It should evaluate the action by showing whether what has been achieved is above or below expectation. It should serve as a basis for correctly planning the next step, (for) modifying the 'overall plan.' Finally, it gives the planners a chance to learn, that is, to gather new general insight . . . regarding the strength or weakness of certain . . . techniques of action. . . .
>
> Rational social management, therefore, proceeds in a spiral of steps each of which is composed of a circle of planning, action, and fact-finding about the result of the action. (pp. 333-334).

This sequential and spiraling method of problem solving is now widely used in organizational development efforts, and has been applied in a variety of industrial, human service, and educational action research projects; and it appears to be at the heart of Tharp and Gallimore's (Note 4) Evaluation Succession model.

Several varieties of activity are often called action research (Chein, Cook, & Harding, 1948). Sometimes the effort is limited to diagnosis and recommendations; sometimes organizations or project implementers carry out the entire process; sometimes records or diaries of actions taken and their perceived effects are maintained. Tharp and Gallimore (Note 4) describe several ways of "knowing," each appropriate to different stages in the development of a program. What they call "experimentation," "qualitative/personal knowing," "data guidance," and "program evaluation" are all useful in program development and research. But the variety of action research most productive of

trustworthy knowledge is experimental action research. Unfortunately, experimental action research is also the most difficult to perform, because it requires the conditions necessary for confident inference (Cook & Campbell, 1979; Gottfredson, Note 1, Chap. 3), and a stable set of interventions that the organization knows how to and can implement in testable form. Seldom do OD efforts aim to implement experimental action research, largely because it is so difficult. Implementing experimental action research is, however, a chief aim of the Program Development Evaluation method. Because of the pace of organizational change, rigorous examination of the consequences of actions may not always be timely and less rigorous ways of knowing about the effects of innovations are often necessary. But research must be coordinate rather than subordinate to problem solving; solving problems without learning how or why they were solved will contribute little to organizational effectiveness or to theory in the long run.

The PDE Method

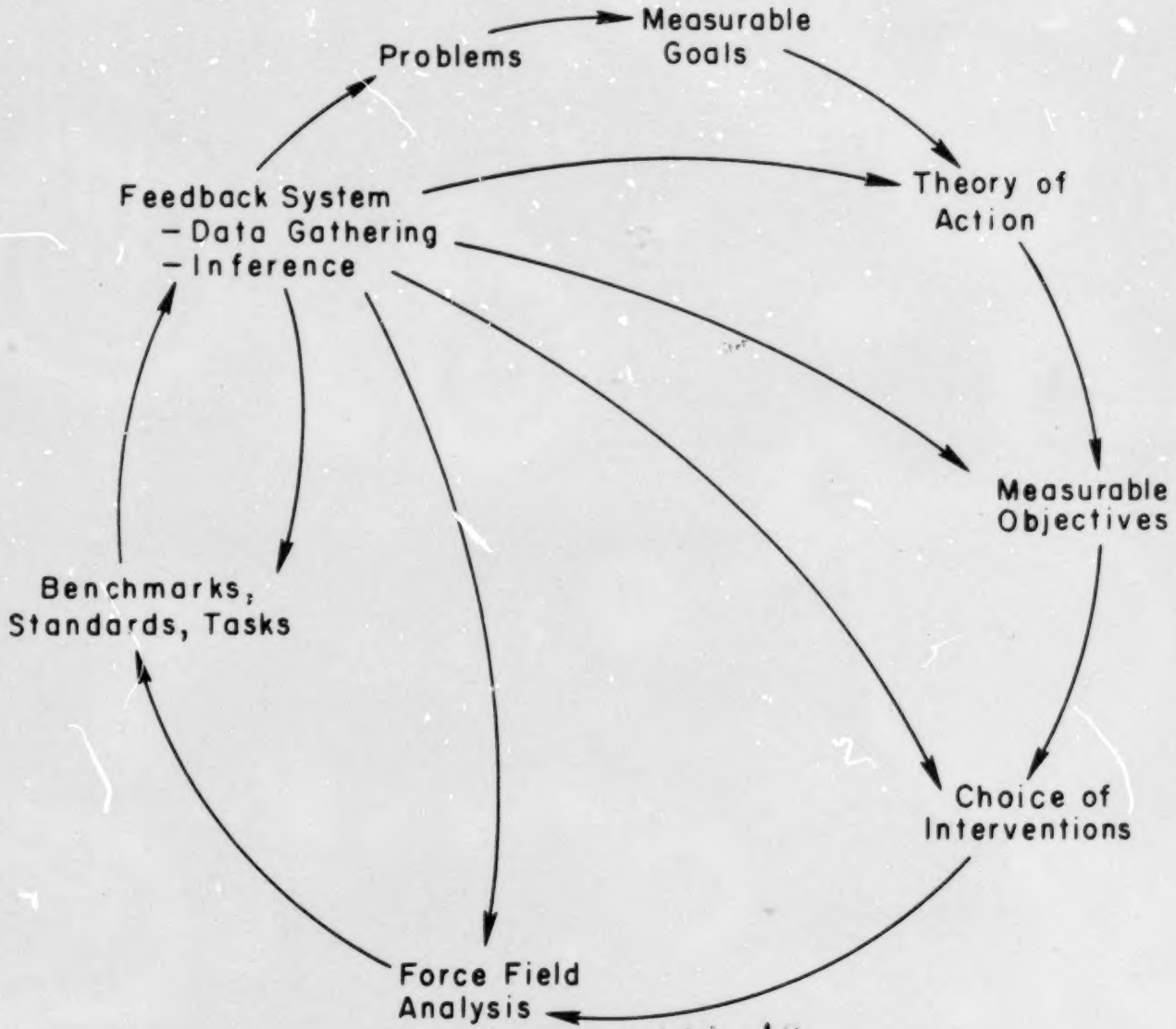
The PDE method emphasizes (a) theory, (b) measurement, and (c) experimental or quasi-experimental design to a greater extent than do many approaches to organizational change and development. In addition, some commonly used terms (most notably "objectives") are defined in a special way. Mastery of the practical meanings of the terms in the PDE structure will provide change agents with a language for thinking about, facilitating, managing, and studying their change efforts.

The Program Development Evaluation method, illustrated in Figure 1, incorporates theory as an explicit component, gives measurable goals and objectives a hard-nosed meaning, and incorporates planning for evaluation implementation in the same way that planning for any other aspect of a project is incorporated. It allows project implementers and evaluators to monitor critical benchmarks in the adoption of any strategy to create change, and it allows them to monitor implementation standards in the implementation of interventions. The principal concepts involved in the PDE structure are elaborated below:

Problems and Goals

Work in any project should begin with an exploration of its intent. Most organizations, and most researchers, have multiple aims. Within the PDE framework, an overarching aim is called a goal. A goal is the obverse of a problem; it specifies how the level of the problem may be measured and therefore how one may know if progress is being made. Several secondary questions are important when discussing goals. The first question serves to reduce ambiguity and enable evaluation; it asks how each goal may be measured. The second question serves to promote realistic research designs; it asks when a project can realistically expect to make a substantial difference and therefore specifies the duration of intervention and the timing of measurement. And the third question, essential in experimental or quasi-experimental action research, asks how one may know that the project itself was responsible for progress towards the goal. These questions are, of course, steps toward involving project implementers in the design of the research. And, they serve to make explicit what the organization expects to accomplish.

Figure 1: The Program Development Evaluation Method



Theory

Actions are taken for reasons that are either articulated or unarticulated. The PDE method is a vehicle for making theory explicit. This is useful because, as the Panel on Research on Rehabilitative Techniques (Martin, Sechrest, & Redner, 1981) notes:

> In attempting to solve any problem, a clear idea of the nature of the problem, its causes, and developmental processes is vital. In the absence of an adequate conceptual framework . . . the rush of enthusiasm for an interesting intervention is likely to short-circuit consideration of these factors. The result is . . . efforts that may be unrelated to the causes of crime, ignore the most suitable target populations, and fail to consider questions of optimal timing and strength of the intervention. The adoption of a theoretical framework necessarily prompts consideration of the above factors and, one hopes, thoughtful development and implementation of . . . interventions, thereby increasing the chances for effectiveness. (p. 29)

Theory helps to organize knowledge and to communicate, it provides a guide for action, and it assists in developing and assessing interventions.

> "Once a basic problem is stated in theoretical terms, planners have an explicit foundation on which to build an intervention strategy and from which to derive a research strategy in conjunction with the intervention" (Martin et al., 1981, p. 34; cf. Glaser, 1980).

In short, an explicit theory provides a template for project implementers' use in building their interventions, as well as a template by which both implementers and researchers can assess those interventions. Therefore, the PDE process calls for deliberate and careful consideration of the question, "Why do these problems exist?" When an intervention is designed using theory, its evaluation tests the theory undergirding the intervention.

Objectives

In the language of PDE, an objective is an intermediary outcome that a project's theory of action implies is important. Like goals, objectives must be stated in measurable terms.

Some examples may help make the distinction between goals and objectives clear. Suppose that a change agent wishes to decrease the death rate due to gastroenteritis in a rural society. The change agent theorizes that the suffering and death are due to the contamination of village water supplies with the cholera microorganism. This theory might suggest a campaign to chlorinate wells, with the objective of decreasing this contamination. The objective would be measured by laboratory analyses of well-water samples to determine the levels of microbial contamination, and attainment of the goal might be measured by counts of deaths per 100,000 population due to gastroenteritis. Another change agent might see the problem somewhat differently. This second change agent may theorize that the suffering and death are due to poor

environmental sanitation: Because few villagers use sanitary latrines, well water is easily contaminated and the cholera microorganism spreads from infected to uninfected persons. This theory might suggest an environmental sanitation campaign directed at persuading villagers to construct sanitary latrines and sanitary wells. The objective now involves villager behavior, and might be measured by the proportion of households using sanitary latrines and water from protected wells. A theory can, of course, suggest multiple interventions and multiple objectives. The second change agent's theory would also reasonably imply chlorination of wells and assessments of well water. The more comprehensive a theory, the more complex the array of interventions and objectives it is likely to suggest.

Change agents could develop theories at many levels to explain the problem of cholera deaths, and each level would suggest somewhat different interventions. To continue the examples, change agents might attribute the problem to (a) normative beliefs in village societies that current standards of environmental sanitation are adequate, (b) the poverty and segregation of the rural people, which deprive them of the resources to build sanitary devices and concentrates them so that they are at high risk, (c) social stratification that allows only an elite merchant class access to sufficient resources to enjoy a sanitary environment, (d) stratification in the world system that enables capitalist countries to keep countries with rural rubber-tapping populations impoverished and the cost of raw materials low. Each of these theories may have considerable validity. Yet each would imply different interventions to solve the problem, ranging from dumping chlorine in wells to overthrowing the capitalist world system. No single cholera prevention project is likely to attempt interventions at all of these levels, and so will not have objectives at each level. A project's theory of action--the theory that drives its interventions--is the theory that is relevant in specifying objectives.

Again, answers to several key questions--how objectives may be measured, when effects are to be expected, and how one may know that the intervention caused the effects--serve to create the evaluation or research design. In addition, the explicit statement of objectives serves to make clear to the organization what it expects to happen as a result of its efforts.

Intervention

An intervention is an action taken to achieve an objective or set of objectives. Ordinarily, it is a major component of a project. The term is often synonymous with "change," "treatment," or "component." Some interventions are aimed at changing the behavior, attitudes, or status of individual people; others are aimed at changing the behavior of an organization or collectivity. An intervention is a process, action, structure, rule, or substance that is applied or put in place to achieve an objective or set of objectives, and therefore to move closer to achieving a goal. An intervention may be chemical, physical, biological, behavioral, social, political, or structural. The interventions employed should be aimed at the objectives an organization's theory of action implies must be achieved.

Forcefield

A forcefield is the social-psychological field that immediately surrounds a decision or action. It includes the forces that compel or restrain against alternative actions as they are perceived by an individual or corporate actor. The notion of a forcefield comes from Lewin's (1951) ideas about the field of forces influencing action. An examination or analysis of an organization's forcefield, especially one that focuses on the field in terms of the resources available and the obstacles to action, is frequently useful for four reasons: (a) By focusing on the organization's perceptions of environmental influences, the nature of these perceptions becomes explicit and open to scrutiny, revision, amendation, supplementation, and test. (b) A complete account of obstacles and resources decreases the likelihood that either pitfalls or potentials will be overlooked in the development of a project. (c) Using knowledge of the influences in the project's environment helps to capitalize on opportunities or arrangements that go beyond the resources under a project's direct control. (d) Alternative strategies or plans to implement any intervention can be created and assessed in the context of the forcefield. Careful attention to the forcefield surrounding a research project increases the likelihood that both interventions and research designs will be implemented as intended.

Because initial analyses of a forcefield may be objectively incorrect, because perceptions change over time, and because the action of a project may alter its forcefield, the dynamic nature of the field is to be expected. A sensible practice, therfore, is to renew forcefield analysis periodically, especially when any strategy being executed on the basis of an initial forcefield analysis is not working well. Breakdowns in experimental design or in the implementation of an intervention would both required renewed analysis.

For more discussion of forcefield analysis and examples of its application see Hersey and Blanchard (1982, pp. 115-119, 269-272). Practical guidance on working with an organization to analyze its forcefield is provided by Blanton & Alley (Note 5, 103-113).

Strategies

Strategies are plans. According to the PDE method, strategies are developed from a forcefield analysis, just as objectives and interventions derive from a theory of action about a problem. Several possible strategies for implementing a research project or one of its component interventions are likely to exist. The task for project implementers and those who are attempting to facilitate strategy development is to create a plan that is perceived as feasible and attractive. If a critical path in some plan is blocked and no way around the obstacle is perceived, the plan is not a good one. Alternative paths that objectively exist but have not been perceived will not be followed. (This point illustrates why thorough and creative forcefield analysis is helpful.) A strategy that appears workable will make use of an organization's resources to overcome the obstacles to implementation and research. Such a strategy may involve (a) moving around an obstacle, (b) decreasing the strength of the forces working against implementation, (c) turn-

ing an obstacle into a resource, or (d) using a strategy in which the obstacle is irrelevant and need not be overcome.

A fully articulated strategy is composed of three kinds of elements: critical benchmarks, implementation standards, and tasks.

Critical benchmarks. A critical benchmark is a key decision, agreement, action, or arrangement necessary to move forward with a plan. A benchmark is much like a gate that must be opened to move along a path. If the gate does not open, progress in executing the strategy is blocked. The locations of these benchmarks (or the nature of them) are made clear in the process of analyzing the forcefield around an intervention. For example, the forcefield analysis about a project's efforts to provide in-service training for teachers might imply that an obstacle lies in teacher unwillingness to participate in training outside of normal working hours, and that a resource is the authority of the deputy superintendent of schools to grant release time and to allocate the funds for substitute instructors. The deputy superintendent's agreement to grant release time and to authorize the expenditure for substitute teachers would then become a critical benchmark. The deputy superintendent is a gatekeeper (Lewin, 1947, p. 333) whose psychology must be examined to learn how to get the gate opened.

Specifying when a critical benchmark is to be accomplished facilitates management. Any strategy will require a temporal or logical sequence of milestones that must be met. In the foregoing example, a failure to accomplish the critical benchmark would signal the need to devise a new strategy for getting the training done, or the need to seek an alternative to training.

Implementation standards. The second part of a strategy for the implementation of both intervention and research are a set of implementation standards. Sound research practice has long involved attention to survey response rates, the measurement of independent and dependent variables, and the procedures used to induce experimental manipulations. Researchers in experimental social psychology and physiology routinely conduct manipulation checks to determine that implementation standards have been met. In research on the effects of dietary supplementation or deficiencies for example, investigators do not assume that an artificial diet has been composed as intended, but perform quantitative analyses on samples of the actual food to determine that the diet consumed was as intended (e.g., Vohra, Gottfredson, & Kratzer, 1968). And, careful researchers on labeling effects perform manipulation checks to see that interpersonal expectancies have been experimentally modified as intended (e.g., Eden & Shani, 1982).

Implementation standards in a prevention research project might include the specification of such intervention characteristics as (a) the skills, knowledge, and numbers of staff; (b) the frequency, duration, and content of interactions of workers with clients, families, or community organizations; (c) the specific actions to be taken in a range of specified situations; and (d) guidelines for the nature and value of reinforcers to be applied for specific kinds of performance. In general, implementation

standards should be sufficiently concrete and specific that they allow for a comparison of what is being done with what is intended.

Tasks. The third part of a strategy is the set of tasks required to execute it. A task statement specifies who will do what by when. Specifying a person to be responsible for executing a particular task, even when a group will be involved, promotes clarity. And specifying when a task is expected to be completed is an additional management tool.

Critical benchmarks, implementation standards, and tasks all serve important functions in project management and worker reinforcement: They serve to guide an organization's efforts, and they provide one kind of objective standard of achievement. A lack of such objective standards "deprives the workers . . . of their legitimate desire for satisfaction on a realistic basis. Under these circumstances, satisfaction or dissatisfaction with {one's} own achievement becomes mainly a question of temperament" (Lewin, 1946, p. 35).

Development

At the very heart of the PDE method is the expectation that project development and research will be an ongoing process, and that the project's environment is dynamic. Only an effete organization is immobile, at equilibrium. Tension, reassessment, review, replanning, and changes in actions taken are the hallmarks of vigorous projects. Consequently, PDE is a cyclical process of action research as progress is made towards achieving goals and objectives (or as goals and objectives are redefined), as new information becomes available, as the environment changes, and as new research questions emerge.

Development occurs largely through the use of information. Information about the achievement or nonachievement of critical benchmarks signals that the forcefield has been usefully understood, or that developmental effort is required to reassess the organization's forcefield. Information that an objective is being achieved signals that an intervention is effective, and information that an objective is not being achieved signals a reconsideration of the appropriateness, strength, or fidelity of the intervention, and prompts new planning. Information that there is progress towards a goal signals that the organization is on the right track. Information that there is no progress towards the goal may signal several things, depending on the pattern of other feedback. If interventions are being implemented as intended and they are achieving their objectives, the theory is called into question. If objectives are not being met, either the theory or integrity of the intervention, or both, should be scrutinized. Success in bringing about elusive objectives and solving serious problems is not to be expected at once. But the PDE structure is intended to provide interim feedback on progress to enable a strengthening of the project. Evidence of project effectiveness lends support to the theory guiding the project.

Evaluation

The PDE structure is intended to facilitate several kinds of evaluation. The explication of a theory of action allows an assessment of its plausibility, and an assessment of the plausibility or strength of the project's planned

interventions in light of the theory.

By tracking the achievement of critical benchmarks, the structure allows assessment of progress towards implementing an intervention. Monitoring implementation standards provides for the assessment of how well interventions are being executed--it provides manipulation checks for both researchers and project managers. These are key elements of formative evaluation and the conduct of research.

The PDE method is also intended to facilitate rigorous summative evaluation--it promotes experimental action research (Chein, Cook, & Harding, 1948) or at least quasi-experimental action research. It repeatedly asks the question, "How do we know the intervention made the difference?" The implementation of an evaluation design is treated in the same way as the implementation of any other intervention. The PDE method assumes that evaluation is an essential component of effective project development and should receive coordinate effort with other aspects of project implementation. Therefore, forcefield analysis is performed for design and data collection issues just as it is for any other project component. Because project implementers are involved in the research design and in the specification of the research questions, their commitment to strong evaluation is expected to increase. And, because the forcefield analysis focuses on the project implementers' own forcefield--their perceptions of the possible--the immediate environment of the evaluation is taken into account when the evaluation is designed, perhaps mitigating some of the resistance to evaluation activities commonly encountered among implementers.

Strength and Fidelity

The PDE method makes possible the assessment of the strength and fidelity of planned interventions through judgments about theoretical plausibility and benchmark monitoring. This assessment can occur in two ways. First, project implementers can assess the consistency of their interventions and objectives with the theory of action underlying their project. That is, a project implementer can determine whether the objectives sought accord with the theory, and whether the interventions planned will plausibly achieve the project's objectives. In short, theory is a template for making judgments about the appropriateness of interventions and objectives that project implementers can use to quality control their own projects. Second, observers of a project, including researchers, can assess its a priori strength by determining whether the planned interventions will plausibly lead to the objectives or goals of the project by assessing them in comparison to state-of-the art theories in the field and the history of similar projects that have been conducted in the past.

As a project develops decisions to adopt an innovation will have been made, and the nature of the events important to the project changes. Early in a project's development, major issues will have to do with ideas for interventions or with strategies for getting an innovation adopted. At later points, important issues will involve the integrity of the intervention's implementation and the assessment of effectiveness. In the early stages of movement toward the adoption of an

innovation, critical benchmarks will involve events related to the decision to adopt a change. Later, implementation standards or quality control checks on the faithfulness of the implementation will increase in importance.

The documentation and assessment of interventions as implemented may involve detailed manuals for the administration of treatments or programs; descriptions of the characteristics of staff and target groups; and accounts of the duration and scheduling of treatments or events, treatment protocols, or proportion of the population served. Therefore, as projects develop, PDE increasingly focuses on the development of manuals to guide service delivery, make diagnoses, and train the staff. Action research projects using this model will develop strategies to monitor staff performance, provide incentive structures to keep performance according to specifications, and the like. The implementation of those strategies is expected to have two consequences: (a) the plans and their execution would increase the integrity of the intervention, and (b) the information generated by the implementation of these plans would describe the integrity of interventions. Manuals developed in the course of carefully implementing a project will allow for subsequent close replications in future research or application.

Some Examples

Two actual examples of projects now underway provide insight into the way the PDE method enables the translation of practitioner ideas into theoretical terms and promotes project development. First, consider an example of a delinqeuncy prevention project involving a peer group intervention. Basically, the project implementers assume that delinquent behavior is supported by a normative subculture. To prevent delinquency the project intended to compose groups of delinquent and non-delinquent youths and to structure group interaction such that delinquent behavior and expressions of beliefs counter to traditional moral beliefs are confronted in the group. Altering beliefs is an important objective. This project's theoretical rational and statement of intervention standards evolved over a period of over two years in which a field worker for the evaluation (Jane St. John) and project implementers (Sonny Luster and William Kottman) used the PDE method. Early research results for this project implied that theoretical intermediary objectives were not always being met. In early results a measure of belief in conventional social rules was negatively effected by the treatment for non-delinquent group participants. This research outcome was used by the project implementers to revise their treatment specifications and to devise ways to monitor group process more closely. Specifically, group leaders were provided additional training, more attention was given to the composition of the groups, and monitoring of group interaction using the Hill (1977; Note 6) interaction matrix was initiated. Although the development and evaluation of this project is still underway, the early negative result has disappeared in subsequent evaluations.

A second actual example is a project currently being conducted by a Southern urban school system. Essentially, this project's theory assumes that assistance and pressure to improve academic

performance, combined with the opportunity to develop attachments to prosocial others, will result in more rewarding experiences in school, greater attachment, and less truancy and delinquent behavior. This theory suggests that students experiencing academic difficulties and who are disruptive in school should receive special attention from project personnel in the form of tutoring and advocacy on behalf of the student with teachers, parents, and others. The evaluation field worker (Denise Gottfredson) and project staff (Doris Coaxum, Barbara Dilligard, Ann Birdseye, and Martha Stuart) used the PDE method. Early implementation monitoring showed that the intervention was not being implemented as intended, with project personnel interacting mostly with students who did not fall within the target group, and interacting only to a limited extent with target students. Clearer implementation standards were specified for the workers to follow, monitoring of these standards increased, and subsequent research results show improved implementation and have begun to show increased academic performance as measured both by standardized test scores and school grades, and some evidence of decreased school disruptiveness and misconduct when randomly equivalent treatment and control students are compared. Some intermediary outcomes, notably attachment to school, were not affected as expected, suggesting that the program can be strengthened further.

Limitations and Potential Criticisms

The PDE method has some limitations, creates some tensions, and is open to criticism. The most important limitations appear to be that it is complex, it is time and expertise intensive, it fails to completely resolve the tensions summative evaluation causes for project implementers, it is an imperfect mechanism for coping with the separate goals of project sponsors and implementing organizations when these are not completely in accord, and it confronts researchers and implementers with tough decisions involving the sacrifice of rigorous research designs in order to achieve some aspect of project implementation.

Time and Talent

The human behavior required to successfully implement the PDE method is complex, and the method's implementation calls for an investment of human resources. Use of the PDE structure calls for high levels of interpersonal competency, tact, patience, communication skill, and understanding of group relations in organizations. In addition, it calls for a thorough understanding of evaluation methods--measurement, social science theory, experimental and quasi-experimental methods, and statistics. Ironically, this combination of competencies are rarely found in one and the same person, suggesting that a team of workers may be required to implement this approach to organizational change. Furthermore, the cyclical or developmental nature of PDE requires constant (or at least frequent periodic) attention, monitoring, updating, and information communication. This taxing process goes beyond the effort typically expended in research.

Some trade-offs are likely to be involved in staffing an action research effort using the PDE

method. Because staff with the requisite skills are hard to find in a typical organization, training will be required. Experience implies that social scientists trained primarily in research methodology, statistics, and theory can successfully implement the method, but that they require additional training in organization development to do so.

Tensions

Tension appears endemic in evaluation research. Too often in the past, evaluation has been used as a tool for canceling a project--even when positive evaluations could not reasonably be expected at an early stage of project development. Tension is also created by the inherently political environment of action projects, and by environments where the successful project does not rock the boat. The development of sound programs usually requires the expenditure of time and money, and often implies the necessity of arrangements that are disruptive.

Tradeoffs and Research Rigor

Program Development Evaluation is value laden. Participation of project implementers is a fundamental principle in the PDE process; pursuit of the goals and objectives of the implementing organization are generally assumed to be desirable (although open to question). Furthermore, an aim of PDE is to develop the implementing organization's capacity to accomplish its goals. Therefore, researchers and implementers collaborate in evaluation design, question formulation, and planning. As a result, researchers extensively intervene in project development--indeed they become a part of the project.

Some evaluators (Perloff, 1979) see this as undesirable in a summative evaluation because it raises questions about the generalizability of the results to situations where researchers are absent. In addition, just as research needs sometimes intrude in project operations by creating new tasks or structural arrangements, the pursuit of a project's programmatic activities can result in compromises in research design. As Deutsch (1968, p. 466) says, "The danger that confronts the research worker in such situations is the possibility that his research design or methodology will be sacrificed to the achievement of the social-action objective."

This "danger" may account in part for the reluctance academic social scientists have shown to participate in action research. This danger seems a small price to pay in exchange for the opportunity action research creates to contribute to the solution of social problems, although the tradeoffs involving evaluation rigor are painful. In short, the PDE method is no panacea for this tough problem.

Complexity

The PDE method is complex. Unfortunately, each component of the method seems at present to be useful and desirable in action reserch efforts. Nevertheless, this complexity suggests that a more streamlined method is appropriate when doing short-term organizational development interventions that do not have a research purpose. Consequently, for many brief organization development interventions, the selective use of those portions of the PDE structure that seem to be the most relevant for the problem at

hand is more appropriate than attempting to use all parts of the method. Researchers should not abandon any part of the method.

Practical Application

Experience implies that one or another part of this process is useful to project implementers in defining their own jobs, in formulating plans, and in clarifying their intentions. Experience also implies that the entire process is sometimes viewed as burdensome. On the whole, this structure seems a clear improvement over some more traditional research methods because it involves implementers in research planning, because it explicitly attempts to build summative evaluation structures based on an organization's forcefield, and because it focuses on goals and objectives of concern to implementers.

My colleagues and I are using the PDE method in the theory-ridden evaluation of delinquency projects conducted by 17 different organizations. In an area where true experiments are almost unknown (Dixon & Wright, 1975; Gottfredson, 1981; Ogawa, Note 7), six of these projects are implementing true experiments. Our opinion is that the quasi-experimental designs of most of the remaining projects are substantially better than would have been the case had project implementers not been involved in an evaluation using the PDE method. Even where the original research design has broken down, project implementers have usually shared with the researchers the task of creating alternative designs and in developing plans to prevent subsequent breakdowns.

The greatest virtues of the PDE method appear to be (a) its ability to make possible theoretical tests through research on action projects, (b) its ability to elaborate clear measurable intermediary outcomes, or objectives, useful in assessing the effectiveness of interventions in theoretical terms, (c) its ability to provide project implementers with the tools to assess their own efforts by measuring interventions against theory and objectives, (d) its ability to generate creative strategies perceived as feasible to implementers based on the divergent and then convergent thinking that takes place in forcefield analysis, (e) its ability to involve project implementers in the research enterprise by engaging them in the specification of measurable goals and objectives, and in the creation of evaluation designs, (f) its ability to provide short-term assessments of progress through the monitoring of critical benchmarks, implementation standards, and tasks, (g) its ability to enable researchers to understand the nature of a project by translating implementer's ideas into a structured language of theoretical research, and (h) its ability to serve as a structure for communication between researchers and practitioners.

PDE makes serious organizational change and rigorous research more attractive to organizations despite its inability to make it truly palatable to all of them. Ideally, practice and research would be merged into a single enterprise in which rigorous theoretical research becomes an integral component of program operation. It is unrealistic to think that most practitioners will ever acquire all of the technical skills required to systematically conduct rigorous research on their activities (just as in is

unrealistic to expect most researchers to become adroit practitioners). In addition, truly rigorous research is not always called for in the development of a project, and not all action projects test theoretically interesting ideas. When rigorous evaluation is called for, however, the PDE structure is helpful. Program Development Evaluation does not successfully resolve many of the sources of tension in merging action with research, but it is progress.

Reference Notes

1. Gottfredson, G. D. The school action effectiveness study: First interim report (Report No. 325). Baltimore: The Johns Hopkins University, Center for Social Organization of Schools, 1982.

2. Wilkins, L. T., & Gottfredson, D. M. Research, demonstration, and social action. Davis, Calif.: National Council on Crime and Delinquency Research Center, 1969.

3. Pearl, A. Quality control in evaluative research of correctional programs. Paper presented at the National Institute on Crime and Delinquency, July 1962.

4. Tharp, R. G., & Gallimore, R. The ecology of program research and development: A model of evaluation succession. Unpublished manuscript. (Available from Roland Tharp, Department of Psychology, University of Hawaii, Honolulu 96822; or Roland Gallimore, Department of Psychiatry and Biobehavioral Sciences, UCLA, Westwood, California 90024.)

5. Blanton, J., & Alley S. Program development: A manual for organizational self-study. (Prepared under contract no. HSM-42-72-143 from NIMH). Nicasio, Calif: Social Action Research Center, 1975.

6. Hill, W. F. Hill interaction matrix (Rev. ed.). Los Angeles: University of Southern California, Youth Study Center, 1965.

7. Ogawa, D. K. Preventing delinquency: The record of accomplishment. In G. D. Gottfredson (Ed.), School action effectiveness study: First interim report (Report No. 325). Baltimore: Johns Hopkins University, Center for Social Organization of Schools, 1982.

References

Chein, I., Cook, S., & Harding, J. The field of action research. American Psychologist, 1948, 3, 43-50.

Cook, T. D., & Campbell, D. T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Deese, J. Psychology as science and art. New York: Harcourt Brace Jovanovich, 1972.

Deutsch, M. Field theory in social psychology. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology (Vol. 1, 2nd ed.). Reading, Mass.: Addison-Wesley, 1968.

Dewey, J. How we think (Rev. ed.). New York: Heath, 1933.

Dixon, M. C., & Wright, W. E. Juvenile delinquency prevention programs: An evaluation of policy related research on the effectiveness of prevention programs. Washington, D.C.: National Science Foundation, 1975.

Eden, D., & Shani, A. B. Pygmalion goes to boot camp: Expectancy, leadership, and trainee performance. Journal of Applied Psychology, 1982, 67, 194-199.

Empey, L. T. Field experimentation in criminal justice: Rationale and design. In M. W. Klein & K. S. Teilmann (Eds.), Handbook of criminal justice evaluation. Beverly Hills, Calif.: Sage, 1980.

French, W. L., & Bell, C. H., Jr. Organization development: Behavioral science interventions for organization improvement (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1978.

Glaser, D. The interplay of theory, issues, policy, and data. In M. W. Klein & K. S. Teilmann (Eds.), Handbook of criminal justice evaluation. Beverly Hills: Sage, 1980.

Gottfredson, G. D. Schooling and delinquency. In S. E. Martin, L. B. Sechrest, & R. Redner (Eds.), New directions in the rehabilitation of criminal offenders. Washington, D.C.: National Academy Press, 1981.

Hall, G. E., & Loucks, S. F. A developmental model for determining whether the treatment is actually implemented. American Educational Research Journal, 1977, 14, 263-276.

Hersey, P., & Blanchard, K. Management of organizational behavior (4th ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1982.

Hill, W. F. Hill interaction matrix (HIM): The conceptual framework, derived rating scales, and an updated bibliography. Small Group Behavior, 1977, 8, 251-268.

Lewin, K. Action research and minority problems. Journal of Social Issues, 1946, 2, 34-46.

Lewin, K. Group decision and social change. In T. M. Newcomb & E. L. Hartley (Eds.), Readings in social psychology. New York: Holt, 1947.

Lewin, K. Field theory in social science. New York: Harper, 1951.

Martin, S. E., Sechrest, L. B., & Redner, R. (Eds.). New directions in the rehabilitation of criminal offenders (Report of the Panel on Research on Rehabilitative Techniques). Washington, D.C.: National Academy Press, 1981.

Perloff, R. (Ed.). Evaluator inventions: Pros and cons. Beverly Hills, Calif.: Sage, 1979.

Quay, H. Three faces of evaluation: What can be expected to work. Criminal Justice and Behavior, 1977, 4, 341-354.

Sarason, S. B. The culture of the school and the problem of change. Boston: Allyn & Bacon, 1971.

Sanford, N. Whatever happened to action research? Journal of Social Issues, 1970, 26, 3-23.

Sechrest, L., White, S. O., & Brown, E. D. (Eds.). The rehabilitation of criminal offenders: Problems and prospects (Report of the Panel on Research on Rehabilitative Techniques). Washington, D.C.: National Academy of Sciences, 1979.

Vohra, P., Gottfredson, G. D., & Kratzer, F. H. The effects of high levels of dietary EDTA, zinc or copper on the mineral contents of some tissue of turkey poults. Poultry Science, 1968, 47, 1334-1343.